

Executive Steering Committee
For A.C.E. Policy II
(ESCAP II)
Report 24

October 12, 2001

ESCAP II: Results of the Person Followup and Evaluation Followup Forms Review

Tamara Adams and Elizabeth A. Krejsa
Decennial Statistical Studies Division and
Planning, Research, and Evaluation Division

U S C E N S U S B U R E A U

Helping You Make Informed Decisions

EXECUTIVE SUMMARY

This study addresses issues that arose in the evaluation followup that were of interest to the Executive Steering Committee on Accuracy and Coverage Evaluation Policy.

What implications do these data have on the adjustment decision?

These data show that there is a net increase in the number of erroneous enumerations of at least 1,454,915. In addition, there were an estimated 15,235,142 persons that were coded as either unresolved or conflicting by the expert matchers. Some small but unknown portion of these cases would also be considered erroneous, thus increasing the 1,454,915 number. This error resulted in a bias in the dual system estimate that caused an overestimate of the net census undercount.

What coding operations have been performed to collect data about erroneous enumerations?

- Production – The Accuracy and Coverage Evaluation first interviewed people in sample clusters to collect P-sample data. We then matched these people to the census. We performed a person followup field interview to determine a final enumeration status of E-sample people who did not match to the P-sample.
- Measurement Error Reinterview – We used the data from the Measurement Error Reinterview to evaluate the production process. During the Measurement Error Reinterview, we coded the evaluation followup questionnaire which collected information about the enumeration status of the production cases in a sample of the production clusters.
- Person Followup/Evaluation Followup Review – We reviewed a sample of person followup and evaluation followup forms in order to ascertain the correct code based on each form, as well as a final code based on the information at hand. We believe that the Person Followup/Evaluation Followup Review produced more accurate results than the original Measurement Error Reinterview.

What codes were assigned during the Person Followup/Evaluation Followup Review?

- Correct – These people were correctly counted in the census. They are determined to have lived within the search area, according to census residence rules on Census Day.
- Erroneous – These people were counted in the census in error. This included people living outside the search area in other residences or group quarters. This also included geocoding errors, discrepant persons, duplicates, and those people with incomplete information for matching and followup.
- Unresolved – People without enough information to accurately code whether they should be counted within the search area.
- Conflicting – These people have different information in production and on the evaluation followup form and clear determination of the final code could not be made.

Were there errors in the identification of erroneous enumerations for the Accuracy and Coverage Evaluation in the Census 2000?

Yes, we observed the following differences in results from production to the Review:

- Correct to Erroneous – The number of production correct enumerations coded as erroneous enumerations decreased from 2,827,414 during the Measurement Error Reinterview to 1,816,315 in the Review. Of these erroneous enumerations, 1,139,407 were from cases where the census and Accuracy and Coverage Evaluation matched during production, most of which were not followed up during production.
- Erroneous to Correct – The number of production erroneous enumerations coded as correct enumerations decreased from 908,385 in the Measurement Error Reinterview to 361,400 in the Review.
- Net Difference in Erroneous Enumeration Coding – The net difference in the Correct Enumeration to Erroneous Enumeration less the Erroneous Enumeration to Correct Enumeration cells decreased from 1,919,029 in the Measurement Error Reinterview to 1,454,915 in the Review. This number represents the erroneous enumerations found by the evaluation followup interview. The change was in part due to an assignment of some cases as conflicting, an increase in unresolved cases attributed to consistent application of matching rules with the census residence rules, and in part due to matching error.
- Unresolved Rate – The number of unresolved people in the Person Followup/Evaluation Followup Review is 12,640,503 (4.8 percent). The number of unresolved people in the Measurement Error Reinterview was 4,559,691 (1.7 percent). This represents an increase in unresolved cases from the Measurement Error Reinterview to the Review.

The Review of the Person Followup and Evaluation Followup forms identified 1,454,915 net erroneous enumerations not found in production. The Measurement Error Reinterview study, using the Evaluation Followup Interview, (Krejsa and Raglin, 2001) identified 1,919,029 erroneous enumerations not found in production.

What effect did the Review have on the correct enumeration rate?

If we vary the correct enumeration probability between 0.5 and 1 for the new unresolved cases and for the conflicting cases, we see that the correct enumeration rate varies between 95.50 (se=0.19) percent and 97.63 (se=0.12) percent. The production correct enumeration rate for this sample is 97.77 (se=0.10) percent. **This difference equated to an estimated 5.9 million (se=0.4 million) net missed erroneous enumerations if we assumed 0.5 percent correct enumeration probabilities for both the conflicting and unresolved cases.** However, the estimated number of missed erroneous enumerations is sensitive to the assumptions that we make about the conflicting and unresolved cases.

1. BACKGROUND

In July, 2001 results from the Measurement Error Reinterview (MER) matching operation using the Evaluation Followup (EFU) forms were presented to the Executive Steering Committee for Accuracy and Coverage Evaluation Policy (ESCAP) committee (Krejsa and Raglin, 2001). Results of this study showed a net increase of approximately 1.9 million erroneous enumerations from the production estimate. This increase seemed large, so additional review of the forms was deemed appropriate. A sample of the Person Followup (PFU) and EFU forms for the E-sample were reviewed to determine the correct code based on each form, the best code, and the reason for assigning the correct code.

1.1 Overview of the PFU Operations

The PFU interviewing phase of the Accuracy and Coverage Evaluation (A.C.E.) involved the followup of persons to resolve match, residence, and enumeration information for the P-sample (Population Sample) and the E-sample (Enumeration Sample). The P-sample consisted of persons from the A.C.E. Person Interview who were listed on the A.C.E. Independent Roster and who were, according to census residence rules, members of the household on Census Day or whose residence status was unresolved. The P-sample was used to estimate missed people in the census. The E-sample comprised the final roster of persons from the census in the A.C.E. sample block clusters. The E-sample was used to estimate erroneous enumerations in the census.

After the A.C.E. housing unit and the person interviewing operations were completed, the person followup matching process was conducted. There were four major steps to the production person matching and followup process:

- Computer Match - The P-sample and the E-sample people were matched by computer. The results were used during the before followup clerical matching.
- Before Followup Matching—The clerical matchers reviewed the P-sample and E-sample persons who were not matched or were possibly matched by the computer, and census cases with insufficient information for matching. The matchers also attempted to identify and code duplicated persons within both the P-sample and the E-sample.
- PFU Interview - Unresolved and/or selected unmatched persons were sent to a field interview. During the interview additional information was obtained to help assign a final match and/or residence status to each person. For the E-sample, nonmatches were sent for a follow-up interview to determine if they were correctly or erroneously enumerated in the block cluster. Certain whole household nonmatches in the P-sample were not sent for a person followup interview. Possible matches were also sent for an interview to resolve their match status.
- After Followup Coding –The information obtained in the PFU interview was used to code the match and/or residence or enumeration status of the persons in question.

1.2 Overview of the EFU and MER Operations

This evaluation process focused on the production interviewing and after followup matching steps in the person matching and followup process. The EFU interview was an expanded PFU interview. The same people followed up in the PFU were interviewed using the EFU form in addition to a sample of people (Krejsa 2000) not interviewed in the PFU (including matches and nonmatches as described above). The EFU form was expanded to ask more detailed questions about other residences a person may have and about movers from a housing unit.

Using this expanded information from the EFU, the MER matching process (structured similarly to the after followup matching step) was conducted to determine residence and enumeration status of the persons in question.

1.3 PFU/EFU Review Sample

The sampling for the PFU/EFU Review was performed using the results of the MER. We selected with certainty all cases where the enumeration status changed between production and MER. The remainder of the cases, where the enumeration status from production and MER were the same, were sampled at varying rates yielding a sample size of 17,522 persons.

2. METHODS

Unlike the production and evaluation coding operations, the Review was an analyst-only operation (that is, no technicians or clerks). Each analyst reviewed a workunit of sampled persons, coding the EFU form independently of the PFU form. During the Review, the EFU and PFU forms in this operation were coded using a consistent coding scheme as follows:

- No EFU Reject Rule –An analyst could not reject an EFU form in the Review. In MER, if the EFU provided insufficient information to code a case then the form was rejected. In the PFU/EFU Review an analyst had to assign a code to each case.
- Consistent Application of Residence Rules – A Census Day address must have been listed on the form if a person was determined to have moved or lived at another residence on April 1st. If the Census Day address was not listed, the followup person was coded unresolved. Due to the design of the evaluation followup form, Census Day addresses were not collected for certain cases. This led to differences in match code results. In addition, the MER used slightly different rules for certain other types of cases.
- Eliminate Coding Error – Only analysts coded cases in the PFU/EFU Review. Since technicians worked on MER without a formal quality assurance (QA), we wanted to eliminate any residual coding error from the technicians. In addition, we wanted to eliminate

any coding error in the PFU remaining after the formal sample-based QA of the clerks and technicians.

After coding each form separately, the analyst indicated which form contained the best code - both, EFU, PFU, or conflicting. We used the following rules to select the form with the best code:

- If either form was unresolved, we chose the other (resolved) form. An exception to this rule was when the unresolved form gave the analyst more information. For example, a mailman in PFU reported that the followup person lived at the address, but a relative of the household reported in EFU that the same followup person had a second home but was not sure where it was located. In this type of case, the EFU form gave more information but not enough to code the person erroneous or correct. We chose the EFU unresolved form over the PFU form.
- If both forms were resolved and on one form the respondent was a proxy and on the other the respondent was a household member, we chose the form with a household member respondent.
- If both forms were resolved and the same type of respondent answered both, we picked the form that gave more information.

In some cases a clear determination of the best code could not be made, these cases were termed ‘conflicting.’ We coded a case conflicting in the following circumstances:

- Contradictory Information from the Same Respondent Type – A case was determined to be conflicting when both forms were completed by the same type of respondent - either both were household respondents or both were similar caliber¹ proxies - who provided contradictory information that resulted in a different enumeration status for the followup person.
- Contradictory Geocoding Information – In addition, if one form indicated through geocoding information that the housing unit was in one place and the other form indicated it was elsewhere, the case was coded as conflicting.

In addition to collecting the match codes, we collected the following data:

- PFU and EFU Proxy Information—Who answered each form and was the person the same

¹Similar caliber proxies were people we considered to have been in the position to have had the same knowledge about the household. Family friends, relatives, or close neighbors were similar caliber proxies; mailmen and apartment managers were of similar caliber. The first category generally provided better data than the latter.

- PFU and EFU Why Codes—Codes to tell us why a form was assigned a given match code

3. LIMITS

Some data in this report were obtained from the EFU. The most significant limitation of the EFU is the nine to ten month time lag between Census Day, April 1, 2000, and when the EFU data were collected in January and February, 2001. People moved during that time period. Given the time lag, people could forget or inaccurately report information. The EFU questionnaire was developed, though, to attempt to minimize such problems by asking questions of the respondent that aid them in recalling the correct information.

Another limitation is that the EFU did not have a full field quality assurance program as did the A.C.E. Person Interview and the PFU.

For both the person followup interview and the evaluation followup interview, there is evidence that the questions were not always read as worded (Keeley 1999 and Krejsa 2001). This may have led to varying responses for questions.

Any standard errors presented in this report are simple jackknife estimates and do not fully capture all phases of A.C.E. sampling.

4. RESULTS

The focus of the results presented here are as follows:

- How do the results of the review compare to the results of the MER and to the production results? Specifically, what is the net difference in erroneous enumerations according to the review in comparison with those identified in MER and those identified in production?
- What is the source of the difference between the MER results and the review?
 - ÷ Is there coding error in the PFU? If so, how much?
 - ÷ Is there coding error in the MER? If so, how much?
- Why are cases coded as unresolved? What percent is due to missing a Census Day address, to the application of residence rules, or to noninterviews?
- Why are cases coded as erroneous that were previously coded as correct enumerations? Why are cases coded as correct that were previously coded as erroneous enumerations?

Throughout this document the following abbreviations are used:

PFU1 = Production Code, which included matches without followup and cases with followup
 PFU2 = Code from the Review of the PFU form
 EFU1 = Code from the MER study

EFU2 = Code from the Review of the EFU form
Best = Code of the chosen form in the Review

4.1 Background Data

The results of the MER produced tables comparing production (PFU1) codes to MER codes (EFU1) below. This table is from the “E-Sample Evaluation Report on Erroneous Enumerations in the Measurement Error Reinterview,” (Krejsa and Raglin, 2001) presented in July, 2001.

We see the following results:

- Correct to Erroneous – There are 2,827,414 production correct enumerations that were coded erroneous enumerations in the MER. This includes 54.3 percent that were matched.
- Erroneous to Correct – There are 908,385 production erroneous enumerations that were coded correct enumerations in the MER.
- Net Difference in Erroneous Enumeration Coding – The net difference between the two cells described above is 1,919,029. This number represents the additional erroneous enumerations found by the MER.
- Unresolved Rate – The unresolved rate following the MER coding was 1.7 percent. The unresolved rate for the production cases in the MER sample was 2.6 percent.

Table 1. Production vs. Measurement Error Reinterview (Evaluation)

Production (PFU1)	Evaluation (EFU1)				Percent of Total
	Correct	Erroneous	Unresolved	Total	
Total Correct Enumerations	247,114,898 (6,503,469)	2,827,414 (223,232)	1,424,770 (254,610)	251,367,081 (6,566,293)	95.8
<i>Matched</i>	218,343,361 (6,768,798)	1,534,800 (182,422)	1,086,812 (224,742)	220,964,973 (6,217,824)	84.2
<i>Nonmatched Correct Enumerations</i>	28,771,537 (1,160,932)	1,292,613 (116,974)	337,957 (56,248)	30,402,108 (1,183,903)	11.5
Erroneous Enumerations	908,385 (99,213)	3,118,191 (202,208)	124,641 (23,343)	4,151,217 (238,893)	1.6
Unresolved	2,873,110 (399,655)	928,719 (117,386)	3,010,280 (203,994)	6,812,110 (489,361)	2.6
Total	250,896,393 (6,581,557)	6,874,324 (363,830)	4,559,691 (353,112)	262,330,408 (6,729,865)	
Percent	95.6	2.6	1.7		100.0

(Source: Krejsa and Raglin, 2001)

4.2 What was the net difference in erroneous enumerations according to the Review in comparison with those identified in MER?

Table 2 shows results comparing productions versus best code from the Review. Important Results from Table 2:

- Correct to Erroneous – The number of production correct enumerations coded as erroneous enumerations decreased from 2,827,414 during the Measurement Error Reinterview to 1,816,315 in the Review. Of these erroneous enumerations, 1,139,407 were from cases where the census and Accuracy and Coverage Evaluation matched during production, most of which were not followed up during production.
- Erroneous to Correct – The number of production erroneous enumerations coded as correct enumerations decreased from 908,385 in the Measurement Error Reinterview to 361,400 in the Review.
- Net Difference in Erroneous Enumeration Coding – The net difference in the Correct Enumeration to Erroneous Enumeration less the Erroneous Enumeration to Correct Enumeration cells decreased from 1,919,029 in the Measurement Error Reinterview to 1,454,915 in the Review. This number represents the erroneous enumerations found by the evaluation followup interview. The change was in part due to an assignment of some cases as conflicting, an increase in unresolved cases attributed to consistent application of matching rules with the census residence rules, and in part due to matching error.

- Unresolved Rate – The number of unresolved people in the Review is 12,640,503 (4.82 percent). The number of unresolved people in the MER was 4,559,691 (1.7 percent). This represents an increase in unresolved cases from the MER to the Review.
- Unresolved and Conflicting Cases – There were an estimated 15,235,142 persons that were coded as either unresolved or conflicting by the expert matchers. Some small but unknown portion of these cases would also be considered erroneous, thus increasing the 1,454,915 number.

The failure of A.C.E. production to find these erroneous enumerations is probably due to differences between the forms. The evaluation followup form incorporated an extensive battery of questions regarding residences other than the sample address. The production A.C.E. person interview and person followup interview were less extensive. They asked only basic questions about residences other than the sample address. Production most often missed erroneous enumerations in group quarters or other residence situations. Future coverage measurement studies will face the problem of how to ask sufficient questions about such other residences without significantly increasing respondent burden.

Table 2. Production vs. Best Code from the PFU/EFU Review

PFU1 Code	Best Code from Second Review				
	Correct	Erroneous	Unresolved	Conflicting	Total
Total Correct Enumerations	238,786,314 (6,297,622)	1,816,315 (189,188)	9,151,011 (770,433)	1,613,442 (231,082)	251,367,081 (6,452,556)
<i>Matched</i>	210,222,189 (5,995,657)	1,139,407 (160,901)	8,763,973 (758,243)	563,514 (184,718)	220,689,083 (6,144,855)
<i>Nonmatched Correct Enumerations</i>	28,564,125 (1,240,041)	676,908 (96,217)	387,038 (105,973)	1,049,928 (136,905)	30,677,998 (1,271,329)
Erroneous	361,400 (46,064)	2,936,887 (199,370)	186,418 (30,064)	666,512 (87,071)	4,151,217 (237,530)
Unresolved	2,529,422 (393,331)	664,929 (67,479)	3,303,074 (226,500)	314,685 (45,382)	6,812,110 (488,029)
Total	241,677,134 (6,358,186)	5,418,131 (299,065)	12,640,503 (843,845)	2,594,639 (258,383)	262,330,408 (6,603,343)
Percent	92.13	2.07	4.82	0.99	100.00

The source of the changes from correct to erroneous include, but are not limited to:

- Coding Error – As we will see later and was demonstrated in the Matching Error Study (Bean 2001), there is a small amount of coding error in production.

- Conflicting – We allowed conflicting cases during the PFU/EFU Review. Some of these cases would have been coded as erroneous enumerations during the MER.
- Increased Unresolved Rate – On the EFU form, Census Day addresses were not collected for certain types of cases leading to differences in match code results. Part of the increase in the unresolved rate is due to the fact that the MER used slightly different rules than production for certain other types of cases.

4.3 Why were cases coded as erroneous in the Review that were previously coded as correct enumerations?

Table 3 details the why codes for cases coded as a correct enumeration in production and as an erroneous enumeration in the PFU/EFU Review. It shows why the erroneous code was used for each person. These are the people in the correct to erroneous cell in Table 2 (n=1,816,315). Note that the conflicting cases are not included here.

Table 3. Production Correct Enumerations to Best Erroneous Enumerations

Reason for Erroneous	# of People	% of Reason Total	Overall %
Fictitious	28,729		1.58
Movers	292,950		16.13
Never Lived Here	65,866		3.63
Address Mixup	47,063		2.59
Birth or Death	39,395		2.17
Other Residence—Interview at First Home	4,711		0.26
<i>Other Residence—Interview at Second Home</i>	<i>423,066</i>	<i>100.00</i>	<i>23.29</i>
Other Home	123,555	29.20	6.80
Joint Custody	73,940	17.48	4.07
Visiting	84,985	20.09	4.68
Other Home for Work	62,352	14.74	3.43
Other Types of Second Residences	78,233	18.49	4.31
Other Residence—Unspecified	103,292		5.69
<i>Group Quarters</i>	<i>614,451</i>	<i>100.00</i>	<i>33.83</i>
Dorm	315,406	51.33	17.37
Military/Shipboard	17,375	2.83	0.96
Nursing Home	152,101	24.75	8.38
Other GQ's	129,570	21.09	7.13
Geocoding	120,530		6.64
Other	76,262		4.20
Total	1,816,315		100.00

As shown above, over half of the erroneous enumerations missed by the production were either at a group quarters (33.83 percent) or at a second home (23.29 percent). Of the people who should have been counted in group quarters missed by production, about half (51.33 percent) of the erroneously enumerated people lived in dormitories on Census Day. Of the total missed erroneous enumerations, 17.37 percent were from dormitories.

4.4 Why were cases coded as correct in the Review that were previously coded as erroneous enumerations?

Table 4 details the why codes for cases coded as an erroneous enumeration in production and as a correct enumeration in the PFU/EFU Review. It shows why the erroneous code was used for each person. These are the people in the erroneous to correct cell in Table 2 (n=361,400). Note that the conflicting cases are not included here.

Table 4. Production Erroneous Enumerations to Best Correct Enumerations

Reason for Erroneous on the PFU Form	# of People	Source of Correct Enumeration		% of Total
		# Due to PFU Coding Error	# Due to Additional Information from EFU	
No Person Followup	41,117	0	41,117	11.38
Lived Here	110,098	110,098	0	30.46
Fictitious	30,564	0	30,564	8.46
Noninterview	17,195	17,195	0	4.76
Lack of Knowledgeable Respondent	23,966	23,966	0	6.63
Movers	48,940	14,715	34,225	13.54
Never Lived Here	5,034	0	5,034	1.39
Address Mixup	8,561	8,561	0	2.37
Birth or Death	1,403	0	1,403	0.39
Other Residence	38,475	12,471	26,004	10.65
Group Quarters	13,622	5,465	8,156	3.77
Geocoding	19,136	7,737	11,398	5.29
Other	3,288	0	3,288	0.91
Total	361,400	200,208	161,191	100.00

The largest category of people coded as correct enumerations in the Review, but were erroneous in production, are those who reported living at the sample address on Census Day (30.46 percent). These people represent coding error in the PFU1 (as shown in Table 2).

In fictitious, noninterview, and lack of knowledgeable respondent cases, the EFU interviewer was able to find a respondent that could answer questions about the household (19.85 percent of the cases coded as erroneous in production and correct in the Review). These cases were

previously coded erroneous because the PFU interviewer could not find a respondent or due to coding error.

4.5 What was the matching error in the PFU?

Table 5, below, shows the portion of changes in enumeration status in Table 2 explained by coding error from production (similar to MES).

Table 5. PFU1 Code vs. PFU2 Code

PFU1 Code	PFU2 Code				Total
	Correct Enumerations	Erroneous Enumerations	Unresolved	No Followup	
Matches	4,217,469	5,890	217,971	216,247,753	220,688,783
Nonmatched Correct Enumerations	30,097,373	205,097	326,515	49,013	30,677,998
Erroneous Enumerations	161,998	3,647,765	247,416	94,038	4,151,217
Unresolved	207,010	409,837	6,111,227	84,036	6,812,110
Total	34,683,850	4,268,589	6,903,129	216,474,840	262,330,408

As we can see above, the PFU coding error in this sample is 0.68 percent ($se=0.07$) (the shaded off-diagonal cells divided by the total). This is similar to the Matching Error Study E-sample gross error rate of 0.62 percent (Bean 2001).

4.6 What was the coding error in the MER?

Table 6 shows the EFU2 code, if coded based on production rules. This shows why the EFU1 unresolved rate was low compared to PFU1.

Table 6. EFU1 Code vs. EFU2 Code

EFU1 Accepted/ Rejected	Evaluation Code	EFU2 Code			Total
		Correct Enumerations	Erroneous Enumerations	Unresolved	
EFU1 Accepted	Matches	203,098,834	258,518	3,628,346	206,985,698
	Nonmatched Correct Enumerations	26,601,666	343,053	1,327,640	28,272,359
	Erroneous Enumerations	628,376	4,123,858	1,557,985	6,310,219
	Unresolved	401,786	300,040	3,343,055	4,044,881
	Total Accepted	230,730,662	5,025,469	9,857,026	245,613,157
EFU1 Rejected	n/a	1,642,609	340,580	14,734,061	16,717,250
Total		232,373,271	5,366,049	24,591,087	262,330,408
% of Total		88.58	2.05	9.38	100.00

As shown above, the coding error for accepted cases in the MER was approximately 3.4 percent (shaded off-diagonal cases divided by the total accepted cases). A portion of the coding error can be explained by the differences between PFU/EFU Review and MER coding rules.

The unresolved rate increased from 1.7 percent (as shown in Table 1) to 9.38 percent of the MER cases.

4.7 Why were cases coded as non-erroneous that were previously coded as erroneous in the EFU1?

Table 7 details the why codes for cases coded as an erroneous enumeration in MER and as a correct or unresolved enumeration in the PFU/EFU Review. It shows why the erroneous code was used for each person.

Table 7. EFU1 Erroneous Enumerations to Best Non-erroneous

Reason for Assigned Code from EFU2	Best Correct (n=770,957)	Best Unresolved (n=937,123)	Percent of Total (n=1,708,080)
Lived Here	386,794	0	22.64
Fictitious	5,786	8,492	0.84
Noninterview	6,575	105,974	6.59
Lack of Knowledgeable Respondent	5,592	42,493	2.82
Movers	76,727	395,882	27.67
Never Lived Here	11,341	41,243	3.08
Address Mixup	30,109	0	1.76
Birth or Death	5,386	0	0.32
Other Residence--First Home	110,463	6,775	6.86
Other Residence--Second Home	42,571	8,068	2.96
Other Residence—Unspecified	63,861	321,506	22.56
Group Quarters	12,139	814	0.76
Geocoding	13,613	5,876	1.14
Percent of Total	45.14	54.86	100.00

People coded as correct enumerations in the Review, but coded erroneous in MER, and who reported living at the sample address (22.64 percent) represent coding error in the MER.

As we can see, the largest category of erroneous people who we later coded as either correct or unresolved, were movers (27.67 percent). The bulk of these people became unresolved (83.77 percent). Since the EFU form did not collect Census Day addresses for mover situations and the change in coding rules used in the Review, we expected this result.

4.8 Why were cases coded unresolved?

Tables 8a and 8b examine the cases that were unresolved in either interview. Recall—the EFU2 unresolved rate was 9.38 percent (Table 6). Table 8a examines the unresolved cases in EFU2.

Table 8a. EFU2 Unresolved Cases

EFU2 Code[†]			
Unresolved Reason	Number Unresolved - Rejected in EFU1 (n=14,734,061)	Number Unresolved - Accepted in EFU1 (n=6,513,971)	Percent of Total Unresolved
No Census Day Address	384,243	900,058	6.04
Noninterview	2,797,690	547,459	15.74
Residence Rules ²	11,526,247	5,049,409	78.01
Other	25,881	17,046	0.20
Percent of Total	69.34	30.66	100.00

[†]Unresolved people from MER who stayed unresolved were not included in this table.

As we can see, the majority (69.34 percent) of the unresolved EFU cases were rejected in the original MER coding. These cases would have been on the diagonal in Table 1 and not part of the unresolved rate. Of those that were not rejected, most (78.23 percent) were unresolved due to residence rules. This included cases where we did not know if a Census Day address was recorded on the form. Since this information was not a required entry during coding, we were not able to determine if a value had been inadvertently not filled. This also included those cases where a respondent did not know if there was a second residence. These cases were incorrectly coded in MER.

Table 8b. PFU2 Unresolved Cases

PFU2 Code		
Unresolved Reason	Number Unresolved	Percent Unresolved
No Census Day Address	74,223	9.37
Noninterview	227,614	28.74
Residence Rules ¹	484,657	61.20
Other	5,409	0.68
Total	791,903	100.00

²A person is unresolved due to residence rules if the respondent did not know if the person in question had a second residence.

As we can see from Table 8b, the reasons for the unresolved rate from the PFU questionnaire closely parallels those of the EFU1 Rejects. Again, most of the unresolved cases (61.20 percent) were due to residence rules.

4.9 What was the correct enumeration rate for the PFU/EFU Review Sample?

Since we do not have a probability of correct enumeration for either conflicting cases or new unresolved cases, if we vary the correct enumeration probability for both, we see the correct enumeration rates below. Standard errors are in parentheses.

Production Correct Enumeration rate for the PFU/EFU Review Sample was **97.77** (se=0.10). This rate included the imputation for the production unresolved cases; therefore, it cannot be directly derived from Table 2.

Table 9. Correct Enumeration Rate Varying the Correct Enumeration Probability

Conflicting Cases Possible Correct Enumeration Probability	Unresolved Cases Possible Correct Enumeration Probability				
	0	0.25	0.5	0.75	1
0	93.37 (0.32)	94.19 (0.26)	95.01 (0.21)	95.83 (0.17)	96.64 (0.16)
0.25	93.62 (0.31)	94.44 (0.25)	95.25 (0.2)	96.07 (0.16)	96.89 (0.14)
0.5	93.86 (0.3)	94.68 (0.24)	95.5 (0.19)	96.32 (0.15)	97.14 (0.13)
0.75	94.11 (0.3)	94.93 (0.24)	95.75 (0.18)	96.57 (0.14)	97.39 (0.12)
1	94.36 (0.3)	95.18 (0.24)	96.0 (0.18)	96.81 (0.14)	97.63 (0.12)

If we vary the correct enumeration probability between 0.5 and 1 for the new unresolved cases and for the conflicting cases, we see that the correct enumeration rate varies between 95.50 (se=0.19) percent and 97.63 (se=0.12) percent. The production correct enumeration rate for this sample is 97.77 (se=0.10) percent. This difference equated to an estimated 5.9 million (se=0.4 million) net missed erroneous enumerations if we assumed 0.5 percent correct enumeration probabilities for both the conflicting and unresolved cases. However, the estimated number of missed erroneous enumerations is sensitive to the assumptions that we make about the conflicting and unresolved cases.

5. CONCLUSIONS

The net number of erroneous enumerations not found by the production operations is 1,454,915. Of these, 62.7 percent were matches, most of which did not go to followup. In addition, there were an estimated 15,235,142 persons that were coded as either unresolved or conflicting by the expert matchers. Some small but unknown portion of these cases would also be considered erroneous, thus increasing the 1,454,915 number. This error resulted in a bias in the dual system estimate that caused an overestimate of the net census undercount. In addition, the unresolved rate has increased from 1.7 percent to 4.8 percent.

The failure of A.C.E. production to find these erroneous enumerations is probably due to differences between the forms. The evaluation followup form incorporated an extensive battery of questions regarding residences other than the sample address. The production A.C.E. person interview and person followup interview were less extensive. They asked only basic questions about residences other than the sample address. Production most often missed erroneous enumerations in group quarters or other residence situations. Future coverage measurement studies will face the problem of how to ask sufficient questions about such other residences without significantly increasing respondent burden.

There was coding error in both production and in the MER. We estimated a 0.68 percent coding error in production and a 3.8 percent coding error in the MER.

6. REFERENCES

Adams, T. and Krejsa, E. (August 2001), "Independent Review of a Sample of the Person Followup and Evaluation Followup Forms", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-PT-24 and Planning, Research and Evaluation Division TXE/2010 Memorandum Series:CM-MER-S-4

Bean, S. (October 2001), "Matching Error in the Census 2000 Accuracy and Coverage Evaluation", ESCAP II, #7.

Childers, D. (January 2001), "Accuracy and Coverage Evaluation: The Design Document", DSSD Census 2000 Procedures and Operations Memorandum Series, Chapter S-DT-1.

Keeley, C. (1999), "Evaluation of the Integrated Coverage Measurement/Post Enumeration Survey Person Followup Questionnaire", Census 2000 Dress Rehearsal Evaluation Results Memorandum Series, Chapter C6.

Krejsa, E. (2000), "Measurement Error Reinterview Sample Selection", Planning, Research, and Evaluation Division TXE/2010 Memorandum Series: CM-MER-S-01.

Krejsa, E. (May 2001), "Observation of Accuracy and Coverage Evaluation - Evaluation Followup Interview and Office Edit January 22-24, 2001", Planning, Research, and Evaluation Division TXE/2010 Memorandum Series: CM-GEN-T-12.

Krejsa, E. and Raglin, D. (October 2001), "Evaluation Results for Changes in A.C.E. Enumeration Status", ESCAP II, #3.

Appendix A

A.1 What were the reasons a case was determined to be conflicting?

Table A1. Production Code of Conflicting Cases

Best Code			
PFU1 Code	Conflicting	Total	Percent of Total
Correct	1,613,442 (231,082)	251,367,081 (6,452,556)	0.64
Erroneous	666,512 (87,071)	4,151,217 (237,530)	16.06
Unresolved	314,685 (45,382)	6,812,110 (488,029)	4.62
Total	2,594,639 (258,383)	262,330,408 (6,603,343)	0.99

As we can see, 16.06 percent of the erroneous production cases were conflicting. Below, we examine the recoded information from the PFU/EFU Review for these conflicting cases.

Table A2. Independent Code for the Conflicting Cases of PFU2 versus EFU2

EFU 2 Code				
PFU2 Code	Correct	Erroneous	Unresolved	Total
Total Correct Enumerations	0	665,794	814,694	1,480,488
<i>Matched</i>	0	516,348	516,164	1,032,512
<i>Nonmatched Correct Enumerations</i>	0	149,446	298,530	447,976
Erroneous	668,417	0	80,692	749,109
Unresolved	312,980	52,062	0	365,042
Total	981,397	717,856	895,386	2,594,639

Table A3. Conflicting Cases

Reason for Erroneous	Percent of People Erroneous to Correct (n=668,417)	Percent of People Correct to Erroneous (n=516,348)
Geocoding	25.60	25.14
Movers	31.08	25.32
Group Quarters	10.57	16.18
Other Residence	23.57	23.62
Other	9.18	9.74
Total	100	100

As expected, the cases with geocoding information and movers were the majority of the conflicting cases. Most of the remainder were group quarters and other residence situations.

The conflicting cases in the Correct to Erroneous column followed the same patterns as shown for the Erroneous to Correct. These data show that there was not a bias in choosing one type of form over the other.

A.2 Who responded to the interviews?

Table A4 shows the respondent status of each interview and gives information as to why a particular interview was chosen. Only cases in which the proxy respondent data were available for both PFU and EFU are included. In some cases the proxy respondent information is not applicable, and therefore was not collected. Such cases include noninterviews and matching errors such as duplicates, insufficient information for matching, and fictitious people. None of these cases are included below.

Table A4. Proxy Information in the PFU Form versus the EFU Form

PFU	Accepted Interview	EFU		Total
		Number of Household Respondents	Number of Proxy Respondents	
Non-proxy		N=26,234,810	N=6,249,918	32,484,728
	Both	22,819,812	2,164,948	
	PFU	1,394,346	3,632,479	
	EFU	945,502	229,262	
	Conflicting	1,075,150	223,229	
Proxy		N=4,735,523	N=9,950,057	14,685,580
	Both	3,011,342	5,528,554	
	PFU	303,473	2,689,374	
	EFU	1,145,380	1,020,507	
	Conflicting	275,328	711,622	
No Followup		N=186,023,645	N=24,450,503	210,474,148
	Both	182,617,305	16,458,065	
	PFU	454,259	1,893,672	
	EFU	2,927,574	6,090,161	
	Conflicting	24,507	8,605	
Total		216,993,978	40,650,478	257,644,456

The EFU interview had a higher household member respondent rate (84 percent) compared to the PFU household member respondent rate (69 percent). We expected this due to the longer field time for the EFU interview.

If there were household member respondent interviews for both surveys, then the enumeration statuses tended to agree (87.0 percent—22,819,812 out of 26,234,810 people). If the enumeration statuses did not agree (for example, either the PFU interview was chosen or the EFU interview was chosen), we chose the household member respondent interview more often. If the EFU form was a proxy, the PFU form was chosen in 94.1 percent of the cases (3,632,479 out of 3,861,741 people); if the PFU form was a proxy, the EFU form was chosen in 79.1 percent of the cases (1,145,380 out of 1,448,853 people). We believe this difference was due to the detail of the EFU interview and the inability of proxy respondents to answer detailed questions about a household.

Proxy information was not recorded when the interview was not completed and where there were matching errors (such as duplicates, insufficient information for matching, and discrepant persons). There are 4,685,953 people (weighted) in such cases where no proxy information was recorded:

- 2,206,541 cases where the information was not recorded on either form
- 889,450 cases where the information was not recorded on the EFU only
- 1,589,962 cases where the information was not recorded on the PFU only